CS 466 Fall 2002
Naomi Nishimura

Probabilistic algorithms notes

Types of algorithms

Monte Carlo algorithms return the right answer with high probability, where for a decision problem the error can be *one-sided* or *two-sided*. One-sided error is when for one of the outputs, the probability of error is zero.

Las Vegas algorithms never return wrong answers; the variation comes in running time.

Probability

A *sample space* is a set of all possible outcomes of an experiment (e.g., heads and tails for a coin flip or numbers 1 through 6 for the toss of a die). An *elementary event* is one possible outcome (e.g. heads or 1). An *event* is a subset of all the possible outcomes (e.g. all even values for the toss of a die).

A *probability measure* is a function Pr mapping a number to each event such that:

1. $0 \leq \mathsf{Pr}[A] \leq 1$;
2. $\mathsf{Pr}[\text{whole space}] = 1$; and
3. for disjoint events, $\mathsf{Pr}[A \cup B] = \mathsf{Pr}[A] + \mathsf{Pr}[B]$.

The *conditional probability* of the probability of an event A given that the outcome is in a subset B, written $\mathsf{Pr}[A|B] = \frac{\mathsf{Pr}[A \cap B]}{\mathsf{Pr}[B]}$.

Events A and B are *independent* if $\mathsf{Pr}[A \cap B] = \mathsf{Pr}[A]\mathsf{Pr}[B]$.

Random variables

A *random variable* X is a function mapping an elementary event to a real number; we define the event $X = x$ to be the set of events s such that $X(s) = x$. A random variable is an *indicator variable* if the value is in $\{0, 1\}$.

The *expectation of* X, $\mathsf{E}[X]$ is defined as $\mathsf{E}[X] = \sum_x x\mathsf{Pr}[X = x]$. The principle of *linearity of expectation* indicates that for any X and Y (not necessarily independent), $\mathsf{E}[X+Y] = \mathsf{E}[X] + \mathsf{E}[Y]$.

The *variance of* X, $\mathsf{Var}[X]$, is defined to be $\mathsf{Var}[X] = \mathsf{E}[(X - \mathsf{E}[X])^2] = \mathsf{E}[X^2] - \mathsf{E}^2[X]$. For independent random variables X and Y, $\mathsf{Var}[X+Y] = \mathsf{Var}[X] + \mathsf{Var}[Y]$. The *standard deviation of* X, σ_X, is the positive square root of the variance.

A few probability distributions

The *uniform distribution* is a distribution in which all events are equally likely.

For *Bernoulli trials* (independent experiments like coin flips) with a probability p of success, if X has value 1 if the result is heads, then $\mathsf{E}[X] = p$ and $\mathsf{Var}[X] = p(1-p)$.

A *geometric distribution* is used to measure the number of trials needed to succeed, such as setting X to be the total number of coin flips needed to obtain heads. When p is the probability of heads, $\mathsf{E}[X] = 1/p$ and $\mathsf{Var}[X] = (1-p)/p^2$.

A *binomial distribution* is used to determine the number of heads in a sequence of n coin flips.

Some useful facts about probability

Boole's inequality states that for a set of events $A_1, A_2, \ldots$, either finite or countably finite and either independent or not, $\mathsf{Pr}[A_1 \cup A_2 \cup \ldots] \leq \mathsf{Pr}[A_1] + \mathsf{Pr}[A_2] + \ldots$.

Markov's inequality states that for X nonnegative, $\mathsf{Pr}[X \geq t] \leq \mathsf{E}[X]/t$ or, equivalently, $\mathsf{Pr}[X \geq t\mathsf{E}[X]] \leq 1/t$.

Chebyshev's inequality states that for any positive real t, $\mathsf{Pr}[|X - \mathsf{E}[X]| \geq t\sigma_X] \leq 1/t^2$.

Some useful mathematical facts

The summation $\sum_{k=0}^{n} x^k$ is a *geometric series* and has value $\frac{x^{n+1}-1}{x-1}$ for any real $x \neq 1$. For an infinite summation with $|x| < 1$, the value is $\frac{1}{1-x}$.

For all real x and n such that $n \geq 1$ and $|x| \leq n$, $(1 + \frac{x}{n})^n \leq e^x$.